DOCUMENT RESUME

ED 043 650 240 TE 499 847

AUTHOR Lincoln, Harry B.
TITLE A Study of Computer Techniques for Music Research. Final Report.
INSTITUTION State Univ. of New York, Binghamton.
SPONS AGENCY Office of Education (DHEW), Washington, D.C. Bureau of Research.
BUREAU NO BR-9-B-101
PUB DATE 30 Jun 70
GRANT OEG-2-9-420101-1067(010)
NOTE 13p.

EDRS PRICE EDRS Price MF-$0.25 HC-$0.75
DESCRIPTORS *Computer Programs, *Data Bases, Data Processing, *Music, Music Theory, *Research

ABSTRACT

Work in three areas comprised this study of computer use in thematic indexing for music research: (1) acquisition, encoding, and keypunching of data--themes of which now number about 50,000 (primarily 16th Century Italian vocal music) and serve as a test base for program development; (2) development of computer programs to process this data; and (3) development of special computer typography for music printing. It is recommended that work in each of these three areas be continued as desirable and feasible in musicological research. (Appended are lists of papers and articles summarizing work on this project, samples of music typography showing incipits in interval sequence order, and a computer print-out showing composer, title, page, source, genre, and serial number.) (JMC)

Final Report

Project No. 9-B-101

Grant No. OEG-2-9-420101-1067 (010)

A STUDY OF COMPUTER TECHNIQUES FOR MUSIC RESEARCH

Harry B. Lincoln
Department of Music
State University of New York at Binghamton
Binghamton, N.Y. 13901

June 30, 1970

The research reported herein was performed pursuant to a grant with the Office of Education, U.S. Department of Health, Education and Welfare. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Office of Education position or policy.

U.S. DEPARTMENT OF
HEALTH, EDUCATION AND WELFARE

Office of Education
Bureau of Research

Acknowledgments

Grateful acknowledgment is made for the encouragement, cooperation and professional help given by Mr. Robert Roberson, Director of the Computer Center at State University of New York at Binghamton. The programmer for this project, Mr. Christian Granger, is also a capable music theorist and composer, and his combination of skills and enthusiasm has been a major asset to the work described in this report.

Table of Contents

SUMMARY

This project comprised the third year's work of a three-year research proposal. (The first year's work was carried out under USOE Project No. 7-8276, Grant No. 1-7-078272-3898 and the second year's work under USOE Project No. 8-B-089, Grant No. OEG 0-8-080089-4581). The original purpose of this research was to develop computer programs, necessary data and procedures for using the computer in a special field of music commonly known as thematic indexing. During the third year's work this purpose has been extended toward a broader study of computer techniques for music research including the development of special computer typography for music printing.

Work in three areas was carried out simultaneously during the third year's work:

1. Acquisition, encoding and keypunching of data. The data bank of themes, now numbering some 50,000, serves as a test base for the programs developed under the project and is available for consultation by other researchers.

2. Development of special programs for information retrieval, pattern matching and organization of data for publication by direct photo copy of computer output. Attempts to develop programs for precise pattern matching of melodies (treating the melody as a wave form) have not yet proved successful.

3. Development of necessary hardware, programs and procedures to permit printing of music on the high-speed computer printer. Special type slugs for the printer were designed and manufactured in a joint effort with the IBM Corporation's Glendale Laboratory in Endicott, New York and are now available to the profession. A sample of music printing is appended to this report as Exhibit A.

An appendix lists five articles and five papers summarizing work under this project during the past three years.

BACKGROUND FOR THIS RESEARCH PROJECT

This project comprised the third year's work of a three-year research proposal. The original purpose of this research was to develop computer programs, necessary data and procedures for using the computer in a special field of music commonly known as thematic indexing. During the third year's work this purpose has been extended toward a broader study of computer techniques for music research including the development of special computer typography for music printing.

A summary of the background work for the three-year project was given in the final report to the first year of funded research under project number 7-8276, USOE Grant No. OEG 1-7-078272-3898. This summary narrated the traditional "hand" methods of developing thematic indices of music, including the preparation of the pilot project which had been carried out to test the feasibility of automating these operations. The final project was carried out using the 4,000 opening themes (known as incipits) from 1,000 four-voice pieces of a 16th century Italian repertory known as the frottola. This music was encoded in the machine readable Ford-Columbia Music Representation along with data on composer, title, sources, etc. A computer program was written which extracted the interval sequence from the representation. (For example if the melody ascends the interval of a second, a +2 is printed; if it descends into a second a -2 is printed, etc.) The interval sequence is calculated through the aeventh interval and these sequences can in turn be arranged in numerical order, thus making it possible to group together melodies with the same contour. The pilot study showed that heretofore unidentified duplications and borrowings in the frottola repertory do indeed exist and proved the feasibility of applying these techniques to much larger repertories.

During the first year of the funded project a large amount of data was acquired, encoded and keypunched. Programs were written to extract greater varieties of information and to permit more effective formats in the computer print-outs. Other musicologists were invited to submit incipits to the data bank which had grown to about 25,000 themes. During this period a Conference on Thematic Indexing was held in Washington, D.C., on November 4, 1967, and was attended by twenty-three musicologists and music librarians with interests in thematic indexing.

During the second year of the funded project there was (1) continued acquisition, encoding and keypunching of data, (2) continued merging of materials from other researchers, with the data bank reaching a total of about 40,000 cards of information, (3) continued comparison of print-outs and pattern matching with original music to determine effectiveness of the research design and (4) work was begun on development of special hardware, programs and procedures to permit printing of music on the high-speed computer printer.

The development of the special music typography, mentioned in (4) above attracted wide interest and as its importance became increasingly evident it has modified the direction of the overall project. In addition to the interest in thematic indexing as such, the project has broadened to a study of computer applications, including the development of new hardware, which can be of value to a wide range of music research interests. These interests include not only musicological work, but also potential applications to libraries, secondary music education and certain phases of music publishing.

RESEARCH PROCEDURES OF THE 1969-1970 PROJECT

As in the first two years of the project several procedures were carried out simultaneously during the 1969-1970 work. These are listed below with indication of changes in emphasis and comparison to earlier work.

1. Continued acquisition, encoding and keypunching of data.
The files of data which have been carefully compiled during the past four years represent the most extensive collection of machine readable information on early music available anywhere. This is not a "closed" file but rather a data bank which will be expanded in the future either by the principal investigator himself or with the use of student help available through the Music Department. Music colleagues from other parts of the country have also contributed to this data bank and it is expected that this collaboration will continue. An example of a welcome interest in collaboration is that of Mr. Paganelli, the music librarian at the Conservatory of Music in Bologna, Italy, who is preparing an extended listing of incipits of anonymous 16th century Italian material held by his library.

During the latter part of this last year's project the principal investigator visited music colleagues in Italy, Hungary and Czechoslovakia, describing the work being carried out under this project and inviting participation in the data bank. He was encouraged by the response to these visits. In addition to the interest of Mr. Paganelli, mentioned above, it appears that material will be forthcoming from Professor Filip at the University of Bratislava, Czechoslovakia. Every effort will be made to get correctly encoded material from these sources in order to enrich the local data bank.

2. Modifications in the search for pattern matchings in the original music.
During the pilot project and the first year of the project considerable time was spent comparing the pattern matchings and other computer citations with the original music in order to test the validity of the original thesis of the project. As the total amount of encoded material expanded this became a tedious and time consuming procedure and it was experience with this routine which was one of the motivations for developing a computer printing of the music itself. Because the computer printing is proving successful we have postponed further comparisons of pattern matchings with original music except in those instances where the matchings seem particularly precise or have certain interesting characteristics.

3. Continued development of necessary hardware programs and procedures to permit printing of music under high speed computer printer.
The development of the means of printing music on the high speed computer printer has taken on increasing interest and importance throughout this project. The computer center financed the development and purchase of a minimum set of type characters (at a cost of $2400) from the IBM Corporation Glendale Laboratory in nearby Endicott, New York. These were delivered in Spring 1969 and programming was begun to permit their use. This programming proved to be quite complex since the computer paper moves in only one direction and thus the program must first develop a two-dimensional array in storage and then properly distribute the printing of the characters as the paper moves through. The program has also been devised so that, for example, there is more space following a whole note then following a half note, etc. This spacing of notation is important to appearance and legibility.

Throughout the year we have experimented with various formats and arrangements of materials using the special type characters. From the Moore Business Forms Company we ordered a supply of 40,000 pages of specially designed preprinted staff paper. Since the computer is "pin driven" (by means of equally spaced holes on the left and right edges of the paper) it is possible to maintain a very accurate registration of type and staff lines for thousands of pages.

Exhibit A, attached, shows a sample page (or a xerography copy thereof) of the most recent printing of music by computer. We have type characters permitting printing of whole notes, upper and lower stemmed half and quarter notes, whole and half rests, flags for smaller note values, dotted notes and bar lines. It is obvious that this is a minimum set of type but it is adequate for a simple melody such as those used in indexing a 16th century music. It will be necessary however to develop more characters in the future to fully utilize the process. In June 1970 consulations were begun with the IBM Glendale Laboratory to investigate the possibility of developing a few more characters, especially good and flat sharp signs, eighth rest and quarter rest.

5. A beginning has been made in the use of the computer typewriter terminal for the inputting of data to the computer. A typewriter is faster than a keypunch and permits on-the-spot corrections and changes more readily. At present the data can be typed using a RAX program which simply stores the data and then permits transference of the data to cards during the night when the computer is available for "batch processing" of materials. When we move to a larger computer (an IBM 360/67) in September, 1970, it will be possible to enter the data directly into the project's own data bank.

RESULTS OF THE PROJECT

1. A very large data bank of early music (primarily sixteenth-century Italian vocal music sources) has been developed during the three years of the project and is available to other researchers as well as serving as test material for computer programs written under this project. About 50,000 themes are stored in the data bank along with composer and title information on over 8,000 compositions. Listing by composer, titles and other formats are available to researchers. Exhibits "B" and "C: show examples of such listings. In Exhibit "B" composers are listed alphabetically with titles listed alphabetically under each composer. Exhibit "C" is a double spaced alphabetical listing of all titles, with single spacing when a title is duplicated. This arrangement permits quick checking of duplicate titles, many of them examples of two settings of the same text by different composers. Other listings which have been generated include arrangements by genre and by source.

2. Effective programs have been developed and are operative to compute interval sequences from incipits and to arrange and format data according to interval sequence. The basic programs for arranging by interval sequence were developed during the first year of the project and examples were appended to the final report of the first year's work. Since that time much effort has gone into developing more sophisticated programs to make possible arrangements of incipits taking into consideration both pitch and rhythm. The basic theoretical idea is to treat melody as a wave form developed on the x and y axes of a graph with rhythm as the x (horizontal) axis and pitch as the y (vertical) axis. Given this mathematical representation of a melody it should be possible to use pattern matching programs to find identical or similar melodies, even when melodies have been transposed or slightly changed. Progress has been made on this phase of the project but to date it is not possible to report a completely satisfactory

result because of programming complexities in translating the Ford-Columbia representation of the music into a graphical representation. Work on this important phase of the project will continue in the coming year.

3. A music typography for the high-speed computer printer has been developed. These special characters are now available from IBM and the following listing indicates the catalog number for each character. Note that some characters require two type slugs.

Slug 836288	Dot, upstem, vertical portion of breve rest	Slug 836291	eighth note flags
Slug 836289	upstem and downstem segments	Slug 836292	Half and whole rests
Slug 836290	upstem and downstem segments eighth flag, upstem	Slug 836293	Black note heads
		Slug 836294	Open note heads

It is expected that in the academic year 1970-1971 several new characters will be added to the above set. The first to be developed will be quarter rest, eighth rest and the flat signs (on line and on space).

CONCLUSIONS AND RECOMMENDATIONS

1. This project has demonstrated the desirability and feasibility of using the computer as a tool in musicological research, with emphasis on its use in thematic indexing. The effort concentrated on three areas: a) acquisition of data, b) development of computer programs to process this data and c) development of special hardware (equipment) components to make possible the printing of music by the high-speed computer printer. It is recommended that the following work be continued in each of these three areas:

a. Acquisition of data. Although the present data bank is large enough for demonstrating the effectiveness of programs as they are developed it is by no means a complete listing of sixteenth-century Italian vocal musical. There is a need in the profession for complete thematic indices of the madrigal, the motet and of other repertories. Expansion of the present listing in any of these areas will necessitate work not only by the present investigator but by other contributing scholars. It may be desirable in the future to formalize this cooperation by setting up an international committee within a professional organization such as the International Musicological Society.

b. Development of computer programs to process data. As indicated earlier in this report there is stil need for more sophisticated programs to order melodic incipits in a manner permitting precise citations of matchings and similarities. The successful solution of this problem may require consultation with persons in mathematics or outside programmers who have had extensive experience in problems of pattern matching.

c. Development of special hardware (equipment) components. At present the emphasis in hardware development in this project lies in the area of music typography. This should be continued until a large enough set of music symbols is available to permit printing of any standard music notation. Because of limited space on the print chain of the computer it will eventually be necessary to purchase a chain devoted only to music research and not shared with other disciplines as at present.

2. It has been increasingly evident to this researcher that there are potential library applications of this research. Although the thematic index is a standard research tool, it is expensive to print in the usual format. The music typography described above makes it possible to print the standard library card (3x5) containing music notation as well as textual information. There are many potential applications, including folk song repertories, popular music, lieder, to name but a few, which deserve investigation.

3. Finally, it is recommended that the computer typewriter terminal be investigated as a means of entering data and of performing certain calculations in this research. A limited amount of work has been done on the Binghamton campus to explore this possibility and with the installation of a new computer in Fall 1970 it is hoped to pursue this further. The present project depends on card input. It is still not clear that replacing card input with typewriter terminal input would result in a clear advantage, but it is a possibility that should be explored. Ultimately, the use of terminal input and output will make the data bank directly accessible (by phone lines) to scholars at distant points.

APPENDIX

Summary listing of papers, articles and reports covering work in using the computer in music research during three years of funded research (from USOE), 1967-1970

Articles

"Musicology and the Computer: The Thematic Index," in Computers in Humanistic Research (E. Bowles, editor), Prentice-Hall, 1967, pp. 184-192.

"Some Criteria for Preparation of Thematic Indices by Computer," in Elektronische Datenverarbeitung in der Musikwissenschaft, (H. Hackmann, editor) Regensburg, 1967, pp. 57-62.

"The Thematic Index: A Computer Application to Musicology" in Computers and the Humanities, Vol. 2, No. 5, May, 1968, pp. 215-220.

"A Computer Application in Musicology: The Thematic Index", Information Processing 68 (International Federation of Information Processing Societies), Amsterdam, 1969, pp. 957-961.

"The Computer and Music Research: Prospects and Problems," in Bulletin of the Council for Research in Music Education, Fall, 1969, pp. 1-9.

Papers

"The Frottole Repertory: A Pilot Study in Information Retrieval," national meeting of American Musicological Society, New Orleans, December, 1967

"A Computer Application to Musicology: The Thematic Index," International Federation of Information Processing Societies, Edinburgh, Scotland, August 15, 1968.

"A Computer Produced Thematic Catalog," New York State Chapter of the Music Library Association, Ithaca, New York, November 2, 1968.

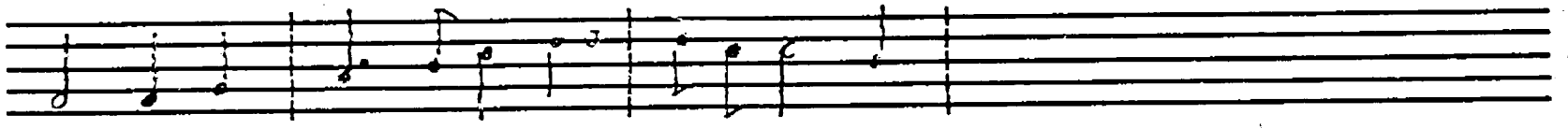
"Recent Developments in Computer Applications to Musicology," Washington, D.C. Chapter of the American Musicological Society, February 18, 1969.

"Teaching Computer Applications in Musicology," Symposium on Introducing the Computer into the Humanities, IBM Department of Education, Poughkeepsie, July 1, 1969.

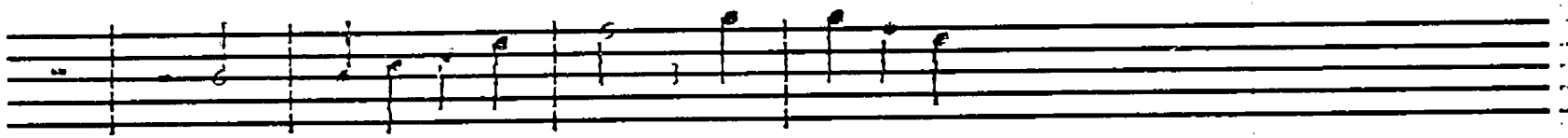
(Editor) The Computer and Music, 354 pp., Cornell University Press, 1970. (Preface by editor)

EXHIBIT "A" MUSIC TYPOGRAPHY SHOWING INCIPITS IN INTERVAL SEQUENCE ORDER

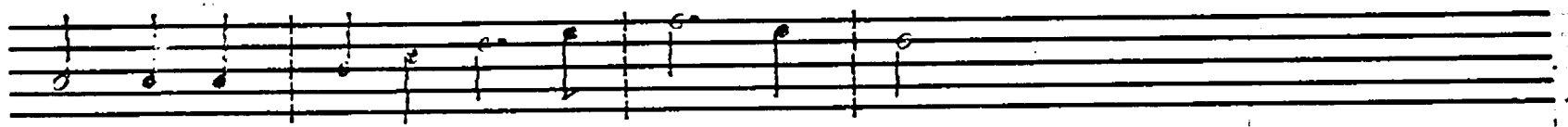
INTERVAL SEQUENCE	CLEF/KEY	SERIAL #	VOICE CODE	TITLE	COMPOSER
+02+02+02+02+02-02-02	|G |M$	011490	0104	MAGN. QUINTI TONI I 4V	GPALEST



+02+02+02+02+02-02-02	|F |K1- |M$	011470	0404	MAGN. PRIMI TONI I 4V	GPALEST



+02+02+02+02+02-02-02	|G |M$	010628	0205	NON BASTA CHəUNA VOLTA	GPALEST



+02+02+02+02+02-02-02	|5C |K1- |M$	007108	0103	U1 RE MI FA SOL LA	AGABRLI

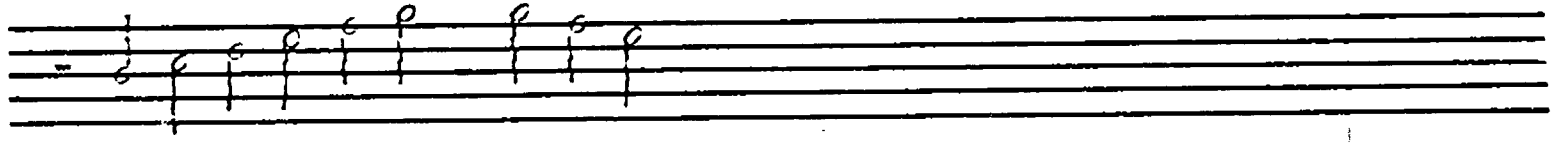


EXHIBIT "A"

EXHIBIT "B"

COMPOSER	TITLE	PAGE	SOURCE	GENRE	SERIAL #
GCROCE	OR CHE SIAM QUI D'INTORNO	345	TORCH2		HL0260438
GCROCE	PENATO IO LUNGAMENTE	EIN1504			HL0070058
GCROCE	S'A LA GELATA MIA TIMIDA	EIN1503			HL0069938
GCROCE	SARA OGNI RE	EIN1504			HL0070138
GCROCE	SIGNOR NEL TUO FUROR	EIN1504			HL0070088
GCROCE	SIGNOR NON MI RIPRENDER	EIN1504			HL0070048
GCROCE	SOCCORIMI SIGNOR	EIN1504			HL0070178
GCROCE	TIRSI MORIR VOLEA	EIN1503			HL0070008
GCROCE	TU SERETO GENTILE	EIN1503			HL0069878
GCROCE	VAGHI E CARI AUGELLINI	EIN1503			HL0069898
GCROCE	VALLI PROFUNDE	EIN1503			HL0069838
GCROCE	VERMIGLIE ROSE	EIN1503			HL0069888
FCRTCIA	CON LEI FUSS'IO DA CHE SI PARTE	36	542/17	MADRIG	HL0409408
FCRTCIA	DONNA VOSTRA BELTA S'A DIR LO	33	542/17	MADRIG	HL0409378
FCRTCIA	NON VED'HOGGI'L MIO SOLE	29	542/17	MADRIG	HL0409338
FCRTCIA	PERCHE LA VITA E BREVE E L'ING	31	542/17	MADRIG	HL0409358
FCRTCIA	VIVACE FIAMMA CHE CON BUBBIA	30	542/17	MADRIG	HL0409348
ADALUA	BEATA NOBIS GAUDIA	P.11	CHWK060	HYMN	HL0404768
ADALUA	CHRISTE REDEMPTOR OMNIUM	P.38	CHWK060	HYMN	HL0404868
ADALUA	TIBI CHRISTE SPLENDOR	P.35	CHWK060	HYMN	HL0404858
ADALUA	UT QUEANT LAXIS	P.22	CHWK060	HYMN	HL0404818
ADALUA	VENI CREATOR SPIRITUS	P.13	CHWK060	HYMN	HL0404778
ADALUA	VEXILLA REGIS PRODEUNT	P.5	CHWK060	HYMN	HL0404738
JDANOLA	CINGARI SIMO	P.13	CHWK043	LIEDER	HL0403328
JDANOLA	MADONNA NOI SAPIMO BEN GIOCARE	P.16	CHWK043	LIEDER	HL0403358
JDANOLA	MEDICI NUI SIAMO	P.12	CHWK043	LIEDER	HL0403318
JDANOLA	NOI TRE MADONNA STAMO	P.15	CHWK043	LIEDER	HL0403348

EXHIBIT "C"

COMPOSER	TITLE	PAGE	SOURCE	GENRE	SERIAL #
JCLEMNS	$CAELESTE BENEFICIUM	P.1	CMM4/15	MOTET	HL020080B
GASOLA	$CANTABANT SANCTI	P.5	A-R VI	MOTET	HL022202B
GASOLA	$CANTATE $DOMINO	P.113	A-R VI	LAUD	HL022223B
FGAFURS	$CASTRA CAELI	P.119	ARCMETR	MOTET	HL009979B
JCLEMNS	$CEST A GRAND TORT	P.135	CMM4/10	CHNSN	HL019940B
PVERDLT	$CHI BUFFA	P.31	541/16	MADRIG	HL041431P
ANON	$CHRISTE, CUNCTORUM	P.12	ARCMETR	MOTET	HL02052.
GWERUKE	$CHRISTI $MATER, AVE	P.1	ARCMETR	MOTET	HL020566B
JARCDLT	$COM'ESSER FUOTE AMORE	P.14	541/16	MADRIG	HL041413B
GDUFAY	$CONDITOR ALME SIDERUM	P.23	MINOR	HYMN	HL003904B
GFESTA	$COSI ESTREMA LA DOGLIA	P.30	541/16	MADRIG	HL041429B
JCLEMNS	$CREAVIT $DEUS HOMINEM	P.73	CMM4/15	MOTET	HL020093B
NGOMBRT	$CUR QUISQUAM CORRADAT	P.104	CMM6/5	MOTET	HL030097B
JCLEMNS	$D'AMY PARFAICT	P.100	CMM4/10	CHNSN	HL019933B
LSENFL	$DAS $GLAUT ZU $SPEYER	P.67	MINOR	LIEDER	HL003911B
MJHAN	$DEH PERCHE NON E IN VOI	P.11	541/16	MADRIG	HL041410B
ANON	$DEH PERCHE NON E IN VOI	P.22	541/16	MADRIG	HL041421B
FGAFURS	$DESCENDI IN HORTUM	P.31	ARCMETR	MOTET	HL009949B
GASOLA	$DEUS CANTICUM NOVUM	P.109	A-R VI	LAUDE	HL022222B
ANON	$DEUS, CREATOR OMNIUM	P.14	ARCMETR	MOTET	HL020524B
GASOLA	$DIES SANCTIFICATUS	P.1	A-R VI	MOTET	HL022201B
MJHAN	$DITIMI O DIUA MIA	P.23	541/16	MADRIG	HL041422B
MJHAN	$DITIMI O DIUA MIA	P.29	541/16	MADRIG	HL041428B
ANON	$DOLCE NEMICA MIA	P.21	541/16	MADRIG	HL041420B
NGOMBRT	$DOMINE, NON SECUNDUM	P.6	CMM6/5	MOTET	HL030082B
PVERDLT	$DONNA CHE DEGGIO FAR	P.24	541/16	MADRIG	HL041423B
ANON	$DONNA CHE DI BELLEZZE	P.26	541/16	MADRIG	HL041425B
GSERGLI	$DONNA, EL PIANTO	P.39	CMM32/2	CANZON	HL023573B